

Electronic signature of DNA nucleotides via transverse transport

Michael Zwolak

Physics Department, California Institute of Technology, Pasadena, California 91125

Massimiliano Di Ventra^[*]

Department of Physics, University of California, San Diego, La Jolla, CA 92093-0319

(Dated: February 2, 2008)

We report theoretical studies of charge transport in single-stranded DNA in the direction perpendicular to the backbone axis. We find that, if the electrodes which sandwich the DNA have the appropriate spatial width, each nucleotide carries a unique signature due to the different electronic and chemical structure of the four bases. This signature is independent of the nearest-neighbor nucleotides. Furthermore, except for the nucleotides with Guanine and Cytosine bases, we find that the difference in conductance of the nucleotides is large for most orientations of the bases with respect to the electrodes. By exploiting these differences it may be possible to sequence single-stranded DNA by scanning its length with conducting probes.

PACS numbers: 73.63.-b, 87.15.-v, 87.64.-t

There has recently been an increased interest in charge transport in DNA, due to both its relevance in physiological reactions and to its potential use in molecular electronics. [1, 2, 3] Previous studies have looked into the effect of the base sequence and structural distortions on charge transport, and the interplay between different transport mechanisms. [1, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12] However, much of the research so far has focused on how charge flows *along* the DNA axis. Very few experimental studies have looked into the transport properties of DNA in the *transverse* direction. For instance, scanning probe techniques have been used to image isolated DNA bases or a few bases in single-stranded DNA [13, 14, 15, 16, 17] but no systematic study is available on the relative ability of each base to carry electrical current.

In this paper, we report on charge transport in single-stranded DNA in the direction *perpendicular* to the backbone axis. Our goal is to determine the relative current carried by each nucleotide in the presence of its neighbors. We envision a possible experimental setup as schematically shown in Fig. 1 where two probes of nanoscale dimensions sandwich the DNA in such a way that one or a few bases are spatially confined within the lateral dimension of the probes. Experimentally, this could be realized by a polynucleotide molecule between two electrodes embedded in a nanopore [18, 19] or by adsorbing the polynucleotide onto a metal surface and using a scanning tunneling microscope (STM) tip to measure the current through a single nucleotide. [14, 15] We therefore seek to identify specific electronic signatures of each nucleotide that would allow us to create a *DNA electronic map*. In particular, we investigate the role the electronic and chemical structures of the bases play in the charge transport, as well as the role of interactions between each base and its nearest neighbors. We indeed find that the current through each individual nucleotide provides a unique electrical signature of each base. The

ratio of the current for different nucleotides gives a measure of the difference of their signatures. These electrical signatures change with orientation of the nucleotides, but are only affected slightly by neighboring nucleotides so long as the electrode lateral dimension is comparable with the average “size” of the bases (i.e., a dimension on the order of 1.0 nm, comparable to the nucleotide spacing of 0.7-0.9 nm [15]). In most cases, the current ratios are quite large when comparing between nucleotides of random orientations, except for the ratio of the current through 2'-deoxyguanosine 5'-monophosphate and the current through 2'-deoxycytidine 5'-monophosphate (i.e., nucleotides with guanine and cytosine bases). Such findings may help create a practical sequencing tool based on the electronic read-out of DNA.

We consider a single strand of DNA, e.g., a polynucleotide, between two electrodes kept at a given bias, (see schematic in Figure 1). The electrodes are represented by 3 layers of about 30 atoms arranged in the [111] surface geometry, which gives an electrode surface spanning about 100 \AA^2 . The electrode-electrode spacing is fixed at 15 \AA . Both the area of the electrode surface and the electrode-electrode distance are such that the largest of the nucleotides (i.e., the one with the A base) can be accommodated between the electrodes. Since we are interested only in the *ratio* between currents of different nucleotides (and we consider only linear response), we employ a simple tight-binding model for the electronic structure of the system with s, p_x, p_y , and p_z orbitals for each carbon, nitrogen, oxygen, and phosphorus atom; and s orbitals for hydrogen and gold. We then use a Green's function approach in order to calculate the current through the whole system (electrodes plus DNA). [20, 21] The Green's function \mathcal{G}_{DNA} of the combined electrode-DNA-electrode system is written as

$$\mathcal{G}_{DNA}(E) = [E\mathcal{S}_{DNA} - \mathcal{H}_{DNA} - \Sigma_t - \Sigma_b]^{-1} \quad (1)$$

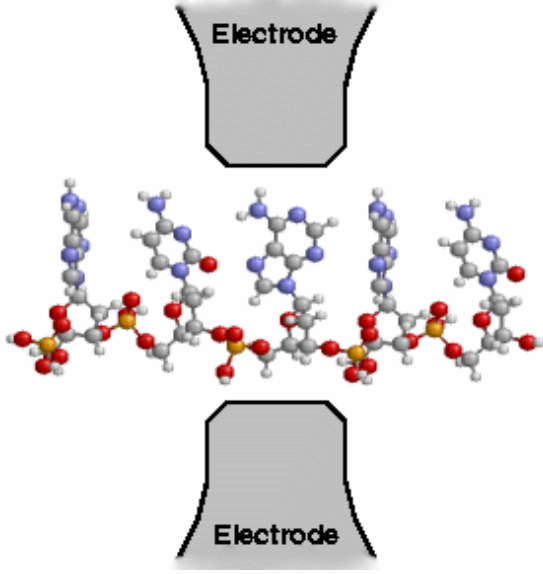


FIG. 1: Schematic of a polynucleotide between two electrodes. The electrodes only couple to a single nucleotide at a time.

where \mathcal{S}_{DNA} and \mathcal{H}_{DNA} are the overlap and the Hamiltonian matrices, respectively. $\Sigma_{t(b)}$ are the self-energy terms that describe the effect of the electronic structure of the leads.

The transmission coefficient $T(E)$ is given by [20]

$$T(E) = T_{tb} = \text{Tr}[\Gamma_t \mathcal{G}_{DNA} \Gamma_b \mathcal{G}_{DNA}^\dagger] \quad (2)$$

where $\Gamma_{t(b)} = i(\Sigma_{t(b)} - \Sigma_{t(b)}^\dagger)$. We calculate the current from

$$I = \frac{2e}{h} \int_{-\infty}^{\infty} dE T(E) [f_t(E) - f_b(E)], \quad (3)$$

where $f_{t(b)} = \{\exp[(E - \mu_{t(b)})/k_B T] + 1\}^{-1}$ is the Fermi function, and $\mu_{t(b)}$ is the chemical potential of the top (bottom) electrode (see Fig. 1). We assume the voltage drops equally in the space between the top electrode and the molecule, and the molecule and bottom electrode. The current is calculated at room temperature. [22]

We first examine the current ratios for the four nucleotides (which we label by A, G, C, or T corresponding to their base) without nearest-neighbor nucleotides attached. The nucleotides are therefore “passivated”, i.e., the oxygen atoms which are normally charged in solution are instead bonded to a hydrogen atom. [1] For each nucleotide, we relaxed the geometry using total-energy Hartree-Fock calculations. [23] We consider first the nucleotides positioned so that the base is in a plane perpendicular to the electrode surface (see Fig. 1). Figure 2 shows the ratio between the current of A and the current of the other nucleotides, I_A/I_X where $X = G, C, \text{ or } T$. The ratios are about 20, 40, and 660 for G, C, and T, respectively. Strong contrast has indeed been observed

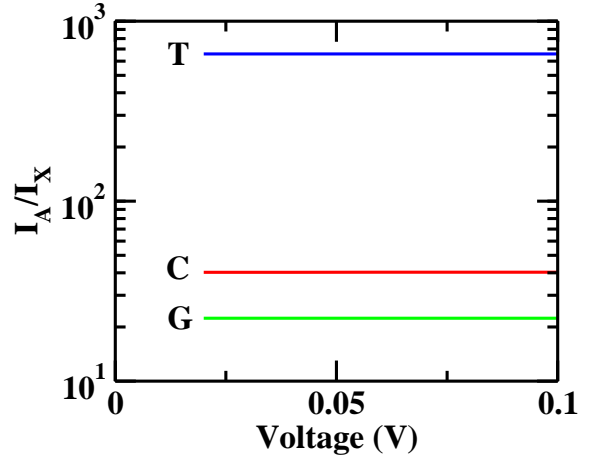


FIG. 2: Current ratios of A to the other nucleotides X ($X = G, C, \text{ or } T$) up to a bias of 0.1V.

in an STM experiment between A and T nucleotides. [14] Figure 2 thus indicates that the isolated nucleotides can be distinguished relatively easily from one another using transverse currents.

The differences in the electrical current ratios are due to the relative position of the Fermi level with respect to the highest-occupied molecular orbital (HOMO) and the lowest-unoccupied molecular orbital (LUMO) as well as the value of the density of states (DOS) [24] of the base at the Fermi level. By projecting onto the different atomic orbitals, we find that the states of the bases make up for most of the HOMO and LUMO levels while the DOS at the Fermi level is mainly determined by the backbone states due to the large coupling of the backbone with the bottom electrode. However, the relative current is closely related to the value of the DOS at the Fermi level due to the bases, which gives the ordering $A > G \sim C > T$. How well those base states couple to *both* electrodes determines the overall magnitude of the relative currents.

Having shown that each nucleotide carries a different electrical signature for a particular geometry, we can now examine how this result changes with orientation of the base with respect to the probe surfaces and location of the backbone atoms relative to the surface. We define our coordinate system as shown in Figure 3. The origin is the nitrogen which attaches each base to the phosphate-sugar backbone. Both rotation and translation involve a rigid movement of the whole nucleotide (base plus backbone). The x-axis runs parallel to the electrode surfaces and perpendicular to the plane of the base. The y-axis runs parallel to the electrode surfaces and in the plane of the base. The z-axis runs perpendicular to the electrode surface. We consider rotations and translations of the DNA nucleotides about each of the axes independently. The six variations we consider are the following: (1) rotation about the x-axis from -10 to +10 degrees, (2) rota-

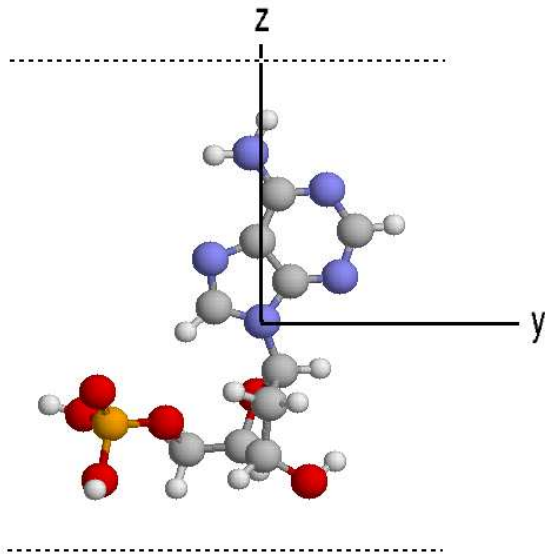


FIG. 3: Coordinate system which we use to define rotations and translations of the nucleotides with respect to the probe surfaces. The x-axis is pointing out of the paper. The dashed lines represent the position of the electrode surfaces.

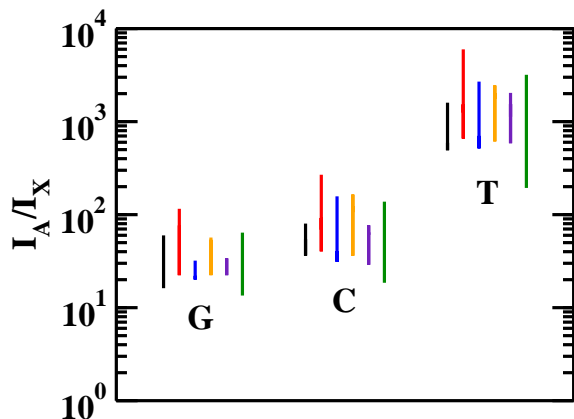


FIG. 4: Current ratios of A to the other nucleotides X ($X = G, C, \text{ or } T$) at a bias of $0.1V$. I_A is for A at its original configuration, while the other nucleotides' orientations are varied. The 6 lines for each nucleotide, from left to right, correspond to rotation about the x-, y-, and z-axes and translation in the x-, y-, and z-directions.

tion about the y-axis from -15 to 15 degrees, (3) rotation about the z-axis from -30 to $+30$ degrees, (4) translations along the x-axis by -1.5 to 1.5 Å, (5) translations along the y-axis by -1.5 to 1.5 Å, and (6) translation along the z-axis by -0.5 to 0.5 Å about the original nucleotide position. [25]

Figure 4 shows relative changes in current with nucleotide rotation and translation as defined above. For each nucleotide, the lines from left to right correspond to

rotation about the x-, y-, and z-axes and to translation along the x-, y-, and z-axes, respectively. The variations (2) and (6) (i.e., rotation about the y-axis and translation along the z-axis, respectively) cause the largest changes in relative current. This is easy to understand since both movements change the distance between the uppermost atoms on the nucleotides and the electrode surface, which gives a large change in coupling. This is analogous to the double-stranded DNA case: variation of the angle and distance between bases can significantly change the conductivity because of a reduction of the inter-base coupling. [12] From Fig. 4, it is clear that, even taking into account the above variations, there is a large distinction between the nucleotides A and T and the pair $\{G, C\}$. The latter two have a similar electrical signature for a wide range of rotations and translations and cannot therefore be easily distinguished. Shot noise, however, being more sensitive to changes in electronic structure than the current, [26] could be used to distinguish between the latter two bases. We leave this study for future investigations.

We conclude by looking at the effect of nearest-neighbor nucleotides on the electrical signature of each individual nucleotide. It is known from ground-state calculations [27, 28, 29] that in isolated stacks of nucleotides the states on each base are highly localized on that base. We therefore expect small corrections due to nearest-neighbor base interactions. We have checked this point by looking at the following representative sequences of three nucleotides of the type AXA, AXT, CXA, GXG, GXT, and TXT, where $X = G, C, \text{ or } T$ is the nucleotide that is sandwiched between the two electrodes. [30] The base of the central nucleotide is also assumed perpendicular to the electrode surface.

The relative change in current for these cases is shown in Figure 5 at $0.1V$. As expected, there is little change in current compared to the individual nucleotides so long as the spatial width of the electrodes is only large enough to accommodate just one base (see Fig. 1). Most of the change is due to the difference in backbone geometry of the strand compared to the isolated nucleotide. Although, for sequences with T as the central nucleotide, the current changes due to nearest neighbors: T carries the smallest current and therefore is the most sensitive to its nearest neighbors interacting with the electrodes. We therefore conclude that the electrode width needs to be about 1 nm if the nearest neighbor bases have uncorrelated orientations. When this is the case, there are variations due to single nucleotide orientation, as shown in Figure 4, but there is not significant variations due nearest neighbor interaction.

In conclusion, we have studied charge transport through nucleotides in the direction transverse to the backbone axis. We find that each nucleotide has a unique electrical signature determined by the electronic and chemical properties of the bases irrespective of the

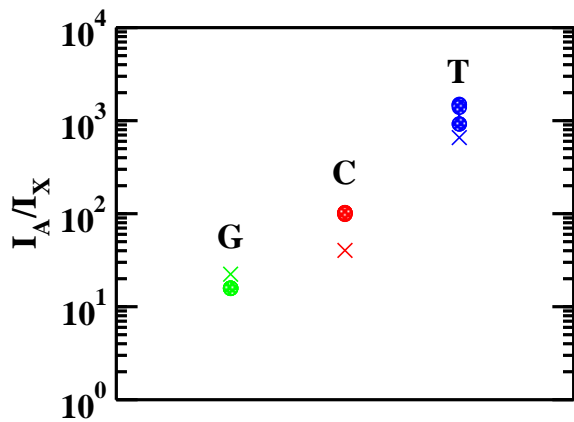


FIG. 5: Current ratios for strands of nucleotides with central nucleotide X ($X = G, C, \text{ or } T$) at a bias of $0.1V$. I_A is for isolated A at its original configuration. The crosses are for the individual nucleotides and the circles are for sequences of three nucleotides (see text for details).

nearest-neighbor nucleotides. Except for the nucleotides with Guanine and Cytosine bases, these signatures are quite robust with variation in nucleotide orientation and location with respect to the electrodes. By exploiting these differences, it may be possible to sequence DNA by scanning its length with a conducting probe. Clearly, for such application many other factors (like solvent and ionic effects) need to be studied in detail and additional theoretical work is necessary. Also, the full noise spectrum may provide extra diagnostic capability to differentiate between bases.

One of us (MD) is indebted to Len Feldman and Mike Ramsey for their suggestion to work on this topic and for useful discussions. We acknowledge partial support from the NSF Grant DMR-01-33075 and the National Human Genome Research Institute. One of us (MZ) also acknowledges support from an NSF Graduate Fellowship.

[*] E-mail address: diventra@physics.ucsd.edu.

- [1] For a recent review, see e.g. Di Ventra, M.; Zwolak, M. DNA Electronics. *Encyclopedia of Nanoscience and Nanotechnology*, Nial, H. S., Eds.; American Scientific Publishers, 2004, Vol. 2, p. 475. See also *Introduction to Nanoscale Science and Technology*, Di Ventra, M., Evoy, S., Heflin, R., Eds.; Kluwer Academics: New York, 2004.
- [2] Boon, E. M.; Livingston, A. L.; Chmiel, N. H.; David, S. S.; Barton, J. K. *Proc. Nat. Acad. Sci. USA* **2003**, *100*, 12543.
- [3] Endres, R. G.; Cox, D. L.; Singh, R. R. P. *Rev. Mod. Phys.* **2004**, *76*, 195.
- [4] Giese, B. *Acc. Chem. Res.* **2000**, *33*, 631.
- [5] Lewis, F. D.; Letsinger, R. L.; Wasielewski, M. R. *Acc. Chem. Res.* **2001**, *34*, 159.
- [6] Schuster, G. B. *Acc. Chem. Res.* **2000**, *33*, 253.

- [7] Li, X.; Yan, Y. *Appl. Phys. Lett.* **2001**, *79*, 2190.
- [8] Berlin, Y. A.; Burin, A. L.; Ratner, M. A. *Chem. Phys.* **2002**, *275*, 61.
- [9] Hjort, M.; Stafstrom, S. *Phys. Rev. Lett.* **2001**, *87*, 228101.
- [10] Zhang, W.; Ulloa, S. E. *Microelectron. J.* **2004**, *35*, 23.
- [11] Roche, S. *Phys. Rev. Lett.* **2003**, *91*, 108101.
- [12] Grozema, F. C.; Siebbeles, L. D. A.; Berlin, Y. A.; Ratner, M. A. *Chem. Phys. Chem.* **2002**, *3*, 536.
- [13] Tao, N. J.; DeRose, J. A.; Lindsay, S. M. *J. Phys. Chem.* **1993**, *97*, 910.
- [14] Tanaka, H.; Kawai, T. *Surf. Sci.* **2003**, *539*, L531.
- [15] Hamai, C.; Tanaka, H.; Kawai, T. *J. Phys. Chem. B* **2000**, *104*, 9894.
- [16] Tanaka, H.; Hamai, C.; Kanno, T.; Kawai, T. *Surf. Sci.* **1999**, *432*, L611.
- [17] Wang, H.D.; Tang, Z.Y.; Li, Z.; Wang, E. K. *Surf. Sci.* **2001**, *480*, L389.
- [18] Akeson, M.; Branton, D.; Kasianowicz, J. J.; Brandin, E.; Deamer, D. W. *Biophys. J.* **1999**, *77*, 3227.
- [19] Li, J.; Gershow, M.; Stein, D.; Brandin, E.; Golovchenko, J. A. *Nat Mater.* **2003**, *2*, 811.
- [20] Datta, S. *Electronic Transport in Mesoscopic Systems*; Cambridge University Press, 1995.
- [21] Rochefort, A.; Di Ventra, M.; Avouris, Ph. *Appl. Phys. Lett.* **2001**, *78*, 2521.
- [22] As in previous calculations in similar structures (see e.g. Di Ventra, M.; Pantelides, S.T.; Lang, N.D. *Phys. Rev. Lett.* **2000**, *84*, 979, and references therein), we implicitly assume that transport is occurring via coherent tunneling, and that coulomb blockade effects are small.
- [23] The relaxed structure of each portion of the backbone which is attached to a base is not a local property of the nucleotide, but is instead a property of the nearby sequence and the environment. We therefore first completely relaxed the 2'-deoxyadenosine 5'-monophosphate, from which we took the relaxed phosphate-sugar structure. We then employed this structure for all other nucleotides. We finally relaxed the latter nucleotides keeping the phosphate-sugar atoms fixed. This also allows each nucleotide to be positioned between the electrodes in a similar way, which ensures that differences in electrical current are due to the bases only.
- [24] Di Ventra, M.; Lang, N.D. *Phys. Rev. B* **2002**, *65*, 045402.
- [25] The choice of the ranges for variations (1) and (6) is dictated by the size of the electrodes, i.e., in an actual experiment and for a polynucleotide, the x-axis rotation and z-axis translation are bound by the distance between probes. The ranges for variations (3), (4) and (5) are limited by the geometry of the [111] surface. The rotation around the y-axis is, on the other hand, arbitrary unless the spatial direction along the x-axis is also limited, as in, e.g., a nanopore.
- [26] Chen, Y.-C.; Di Ventra, M. *Phys. Rev. B* **2003**, *67*, 153304.
- [27] Ladik, J.; Suhai, S. *Phys. Lett. A* **1980**, *77*, 25.
- [28] Otto, P.; Clementi, E.; Ladik, J. *J. Chem. Phys.* **1983**, *78*, 4547.
- [29] Clementi, E.; Corongiu, G. *Int. J. Quant. Chem.* **1982**, *QBS9*, 213.
- [30] The coordinates of these sequences have been relaxed by total-energy Hartree-Fock calculations in a similar way to the isolated nucleotides (see Ref. 23).